Michael R. Berthold
Ad Feelders
Georg Krempl (Eds.)

LNCS 12080

Advances in Intelligent Data Analysis XVIII

18th International Symposium on Intelligent Data Analysis, IDA 2020
Konstanz, Germany, April 27–29, 2020
Proceedings

Springer Open

EXTRAS ONLINE

Lecture Notes in Computer Science 12080

Founding Editors

Gerhard Goos
Karlsruhe Institute of Technology, Karlsruhe, Germany
Juris Hartmanis
Cornell University, Ithaca, NY, USA

Editorial Board Members

Elisa Bertino
Purdue University, West Lafayette, IN, USA
Wen Gao
Peking University, Beijing, China
Bernhard Steffen
TU Dortmund University, Dortmund, Germany
Gerhard Woeginger
RWTH Aachen, Aachen, Germany
Moti Yung
Columbia University, New York, NY, USA

More information about this series at http://www.springer.com/series/7409

Michael R. Berthold · Ad Feelders ·
Georg Krempl (Eds.)

Advances in Intelligent Data Analysis XVIII

18th International Symposium on Intelligent Data Analysis, IDA 2020
Konstanz, Germany, April 27–29, 2020
Proceedings

Springer Open

Editors
Michael R. Berthold
University of Konstanz
Konstanz, Germany

Ad Feelders
Utrecht University
Utrecht, The Netherlands

Georg Krempl
Utrecht University
Utrecht, The Netherlands

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Computer Science
ISBN 978-3-030-44583-6 ISBN 978-3-030-44584-3S (eBook)
https://doi.org/10.1007/978-3-030-44584-3

LNCS Sublibrary: SL3 – Information Systems and Applications, incl. Internet/Web, and HCI

© The Editor(s) (if applicable) and The Author(s) 2020. This book is an open access publication.
Open Access This book is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.
The images or other third party material in this book are included in the book's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

We are proud to present the proceedings of the 18th International Symposium on Intelligent Data Analysis (IDA 2020), which was held during April 27–29, 2020, in Konstanz, Germany. The first symposium of this series was organized in 1995 and held biannually until 2009, when the conference switched to being held annually. Following demand expressed by the IDA community in a survey held in 2018, IDA 2020 was the first of the series to take place in spring rather than fall, as was common before.

The switch to April, and a more organized outreach to the community, coincided with an increase in the number of submissions from 65 in 2018, to 114 in 2020. After a rigorous review process, 45 of these 114 submissions were accepted for presentation. Almost all submissions were reviewed by at least three Program Committee (PC) members (only two papers had two reviews) and a substantial number of submissions received more than three reviews. In addition to the PC, the review process also involved program chair advisors – a select set of senior researchers with a multi-year involvement in the IDA symposium series. Whenever a program chair advisor flagged a paper with an informed, thoughtful, positive review due to the paper presenting a particularly interesting and novel idea, the paper was accepted irrespective of the other reviews. Each accepted paper was offered a slot for either oral presentation (15 papers) or poster presentation (30 papers).

We wish to express our gratitude to the authors of all submitted papers for their high-quality contributions; to the PC members and additional reviewers for their efforts in reviewing, discussing, and commenting on all submitted papers; to the program chair advisors for their active involvement; and to the IDA council for their ongoing guidance and support. Many people have helped behind the scenes to make IDA 2020 possible, but this year we are particularly grateful to our publicity chairs who helped spread the word: Daniela Gawehns and Hugo Manuel Proença!

February 2020

Georg Krempl
Ad Feelders
Michael R. Berthold

Organization

Program Chairs

Georg Krempl	Utrecht University, The Netherlands
Ad Feelders	Utrecht University, The Netherlands

Program Chair Advisors

Niall Adams	Imperial College London, UK
Michael R. Berthold	University of Konstanz, Germany
Hendrik Blockeel	Katholieke Universiteit Leuven, Belgium
Elizabeth Bradley	University of Colorado Boulder, USA
Tijl De Bie	Ghent University, Belgium
Wouter Duivesteijn	Eindhoven University of Technology, The Netherlands
Elisa Fromont	Université de Rennes 1, France
Johannes Fürnkranz	Johannes Kepler University Linz, Austria
Jaakko Hollmén	Aalto University, Finland
Frank Höppner	Ostfalia University of Applied Sciences, Germany
Frank Klawonn	Ostfalia University of Applied Sciences, Germany
Arno Knobbe	Leiden University, The Netherlands
Rudolf Kruse	University of Magdeburg, Germany
Nada Lavrač	Jozef Stefan Institute, Slovenia
Matthijs van Leeuwen	Leiden University, The Netherlands
Xiaohui Liu	Brunel University, UK
Panagiotis Papapetrou	Stockholm University, Sweden
Arno Siebes	Utrecht University, The Netherlands
Stephen Swift	Brunel University, UK
Hannu Toivonen	University of Helsinki, Finland
Allan Tucker	Brunel University, UK
Albrecht Zimmermann	Université Caen Normandie, France

Program Committee

Fabrizio Angiulli	DEIS, University of Calabria, Italy
Martin Atzmueller	Tilburg University, The Netherlands
José Luis Balcázar	Universitat Politècnica de Catalunya, Spain
Giacomo Boracchi	Politecnico di Milano, Italy
Christian Borgelt	Universität Salzburg, Austria
Henrik Boström	KTH Royal Institute of Technology, Sweden
Paula Brito	University of Porto, Portugal
Dariusz Brzezinski	Poznań University of Technology, Poland
José Del Campo-Ávila	Universidad de Málaga, Spain

Cassio de Campos	Eindhoven University of Technology, The Netherlands
Andre de Carvalho	University of São Paulo, Brazil
Paulo Cortez	University of Minho, Portugal
Bruno Cremilleux	Université de Caen Normandie, France
Brett Drury	LIAAD-INESC-TEC, Portugal
Saso Dzeroski	Jozef Stefan Institute, Slovenia
Nuno Escudeiro	Instituto Superior de Engenharia do Porto, Portugal
Douglas Fisher	Vanderbilt University, USA
Joao Gama	University of Porto, Portugal
Lawrence Hall	University of South Florida, USA
Barbara Hammer	Bielefeld University, Germany
Martin Holena	Institute of Computer Science, Czech Republic
Tomas Horvath	Eötvös Loránd University, Hungary
Francois Jacquenet	Laboratoire Hubert Curien, France
Baptiste Jeudy	Laboratoire Hubert Curien, France
Ulf Johansson	Jönköping University, Sweden
Alipio M. Jorge	University of Porto, Portugal
Irena Koprinska	The University of Sydney, Australia
Daniel Kottke	University of Kassel, Germany
Petra Kralj Novak	Jozef Stefan Institute, Slovenia
Mark Last	Ben-Gurion University of the Negev, Israel
Niklas Lavesson	Jönköping University, Sweden
Daniel Lawson	University of Bristol, UK
Jefrey Lijffijt	Ghent University, Belgium
Ling Luo	The University of Melbourne, Australia
George Magoulas	Birkbeck University of London, UK
Vlado Menkovski	Eindhoven University of Technology, The Netherlands
Vera Migueis	University of Porto, Portugal
Decebal Constantin Mocanu	Eindhoven University of Technology, The Netherlands
Emilie Morvant	University of Saint-Etienne, LaHC, France
Mohamed Nadif	Paris Descartes University, France
Siegfried Nijssen	Université Catholique de Louvain, Belgium
Andreas Nuernberger	Otto-von-Guericke University of Magdeburg, Germany
Kaustubh Raosaheb Patil	Massachusetts Institute of Technology, USA
Mykola Pechenizkiy	Eindhoven University of Technology, The Netherlands
Jose-Maria Pena	Universidad Politécnica de Madrid, Spain
Ruggero G. Pensa	University of Torino, Italy
Marc Plantevit	LIRIS, Université Claude Bernard Lyon 1, France
Lubos Popelinsky	Masaryk University, Czech Republic
Eric Postma	Tilburg University, The Netherlands
Miguel A. Prada	Universidad de Leon, Spain
Ronaldo Prati	Universidade Federal do ABC, UFABC, Brazil
Peter van der Putten	Leiden University and Pegasystems, The Netherlands
Jesse Read	École Polytechnique, France
Antonio Salmeron	University of Almería, Spain
Vítor Santos Costa	University of Porto, Portugal

Christin Seifert	University of Twente, The Netherlands
Roberta Siciliano	University of Naples Federico II, Italy
Jerzy Stefanowski	Poznań University of Technology, Poland
Frank Takes	Leiden University and University of Amsterdam, The Netherlands
Maguelonne Teisseire	Irstea, UMR Tetis, France
Ljupco Todorovski	University of Ljubljana, Slovenia
Melissa Turcotte	LANL, USA
Cor Veenman	Netherlands Forensic Institute, The Netherlands
Veronica Vinciotti	Brunel University, UK
Filip Zelezny	Czech Technical University, Czech Republic
Leishi Zhang	Middlesex University, UK

Contents

Multivariate Time Series as Images: Imputation Using Convolutional Denoising Autoencoder 1
Abdullah Al Safi, Christian Beyer, Vishnu Unnikrishnan, and Myra Spiliopoulou

Dual Sequential Variational Autoencoders for Fraud Detection 14
Ayman Alazizi, Amaury Habrard, François Jacquenet, Liyun He-Guelton, and Frédéric Oblé

A Principled Approach to Analyze Expressiveness and Accuracy of Graph Neural Networks 27
Asma Atamna, Nataliya Sokolovska, and Jean-Claude Crivello

Efficient Batch-Incremental Classification Using UMAP for Evolving Data Streams 40
Maroua Bahri, Bernhard Pfahringer, Albert Bifet, and Silviu Maniu

GraphMDL: Graph Pattern Selection Based on Minimum Description Length 54
Francesco Bariatti, Peggy Cellier, and Sébastien Ferré

Towards Content Sensitivity Analysis 67
Elena Battaglia, Livio Bioglio, and Ruggero G. Pensa

Gibbs Sampling Subjectively Interesting Tiles. 80
Anes Bendimerad, Jefrey Lijffijt, Marc Plantevit, Céline Robardet, and Tijl De Bie

Even Faster Exact k-Means Clustering 93
Christian Borgelt

Ising-Based Consensus Clustering on Specialized Hardware 106
Eldan Cohen, Avradip Mandal, Hayato Ushijima-Mwesigwa, and Arnab Roy

Transfer Learning by Learning Projections from Target to Source 119
Antoine Cornuéjols, Pierre-Alexandre Murena, and Raphaël Olivier

Computing Vertex-Vertex Dissimilarities Using Random Trees: Application to Clustering in Graphs. 132
Kevin Dalleau, Miguel Couceiro, and Malika Smail-Tabbone

Evaluation of CNN Performance in Semantically Relevant Latent Spaces 145
Jeroen van Doorenmalen and Vlado Menkovski

Vouw: Geometric Pattern Mining Using the MDL Principle 158
Micky Faas and Matthijs van Leeuwen

A Consensus Approach to Improve NMF Document Clustering 171
Mickael Febrissy and Mohamed Nadif

Discriminative Bias for Learning Probabilistic Sentential Decision Diagrams 184
Laura Isabel Galindez Olascoaga, Wannes Meert, Nimish Shah, Guy Van den Broeck, and Marian Verhelst

Widening for MDL-Based Retail Signature Discovery 197
Clément Gautrais, Peggy Cellier, Matthijs van Leeuwen, and Alexandre Termier

Addressing the Resolution Limit and the Field of View Limit in Community Mining 210
Shiva Zamani Gharaghooshi, Osmar R. Zaïane, Christine Largeron, Mohammadmahdi Zafarmand, and Chang Liu

Estimating Uncertainty in Deep Learning for Reporting Confidence: An Application on Cell Type Prediction in Testes Based on Proteomics 223
Biraja Ghoshal, Cecilia Lindskog, and Allan Tucker

Adversarial Attacks Hidden in Plain Sight 235
Jan Philip Göpfert, André Artelt, Heiko Wersing, and Barbara Hammer

Enriched Weisfeiler-Lehman Kernel for Improved Graph Clustering of Source Code 248
Frank Höppner and Maximilian Jahnke

Overlapping Hierarchical Clustering (OHC) 261
Ian Jeantet, Zoltán Miklós, and David Gross-Amblard

Digital Footprints of International Migration on Twitter 274
Jisu Kim, Alina Sîrbu, Fosca Giannotti, and Lorenzo Gabrielli

Percolation-Based Detection of Anomalous Subgraphs in Complex Networks 287
Corentin Larroche, Johan Mazel, and Stephan Clémençon

A Late-Fusion Approach to Community Detection in Attributed Networks 300
Chang Liu, Christine Largeron, Osmar R. Zaïane, and Shiva Zamani Gharaghooshi

Reconciling Predictions in the Regression Setting: An Application to Bus Travel Time Prediction 313
João Mendes-Moreira and Mitra Baratchi

A Distribution Dependent and Independent Complexity Analysis of Manifold Regularization 326
Alexander Mey, Tom Julian Viering, and Marco Loog

Actionable Subgroup Discovery and Urban Farm Optimization 339
Alexandre Millot, Romain Mathonat, Rémy Cazabet, and Jean-François Boulicaut

AVATAR - Machine Learning Pipeline Evaluation Using Surrogate Model 352
Tien-Dung Nguyen, Tomasz Maszczyk, Katarzyna Musial, Marc-André Zöller, and Bogdan Gabrys

Detection of Derivative Discontinuities in Observational Data 366
Dimitar Ninevski and Paul O'Leary

Improving Prediction with Causal Probabilistic Variables 379
Ana Rita Nogueira, João Gama, and Carlos Abreu Ferreira

DO-U-Net for Segmentation and Counting: Applications to Satellite and Medical Images 391
Toyah Overton and Allan Tucker

Enhanced Word Embeddings for Anorexia Nervosa Detection on Social Media 404
Diana Ramírez-Cifuentes, Christine Largeron, Julien Tissier, Ana Freire, and Ricardo Baeza-Yates

Event Recognition Based on Classification of Generated Image Captions 418
Andrey V. Savchenko and Evgeniy V. Miasnikov

Human-to-AI Coach: Improving Human Inputs to AI Systems 431
Johannes Schneider

Aleatoric and Epistemic Uncertainty with Random Forests 444
Mohammad Hossein Shaker and Eyke Hüllermeier

Master Your Metrics with Calibration 457
Wissam Siblini, Jordan Fréry, Liyun He-Guelton, Frédéric Oblé, and Yi-Qing Wang

Supervised Phrase-Boundary Embeddings 470
Manni Singh, David Weston, and Mark Levene

Predicting Remaining Useful Life with Similarity-Based Priors 483
Youri Soons, Remco Dijkman, Maurice Jilderda, and Wouter Duivesteijn

Orometric Methods in Bounded Metric Data. 496
Maximilian Stubbemann, Tom Hanika, and Gerd Stumme

Interpretable Neuron Structuring with Graph Spectral Regularization 509
Alexander Tong, David van Dijk, Jay S. Stanley III, Matthew Amodio, Kristina Yim, Rebecca Muhle, James Noonan, Guy Wolf, and Smita Krishnaswamy

Comparing the Preservation of Network Properties by Graph Embeddings . . . 522
Rémi Vaudaine, Rémy Cazabet, and Christine Largeron

Making Learners (More) Monotone 535
Tom Julian Viering, Alexander Mey, and Marco Loog

Combining Machine Learning and Simulation to a Hybrid Modelling Approach: Current and Future Directions 548
Laura von Rueden, Sebastian Mayer, Rafet Sifa, Christian Bauckhage, and Jochen Garcke

LiBRe: Label-Wise Selection of Base Learners in Binary Relevance for Multi-label Classification 561
Marcel Wever, Alexander Tornede, Felix Mohr, and Eyke Hüllermeier

Angle-Based Crowding Degree Estimation for Many-Objective Optimization 574
Yani Xue, Miqing Li, and Xiaohui Liu

Author Index 587

DO-U-Net for Segmentation and Counting

Applications to Satellite and Medical Images

Toyah Overton[1,2(✉)] and Allan Tucker[1]

[1] Department of Computer Science, Brunel University London, Uxbridge, UK
{toyah.overton,allan.tucker}@brunel.ac.uk
[2] Alcis Holdings Ltd., Guildford, UK

Abstract. Many image analysis tasks involve the automatic segmentation and counting of objects with specific characteristics. However, we find that current approaches look to either segment objects or count them through bounding boxes, and those methodologies that both segment and count struggle with co-located and overlapping objects. This restricts our capabilities when, for example, we require the area covered by particular objects as well as the number of those objects present, especially when we have a large amount of images to obtain this information for. In this paper, we address this by proposing a Dual-Output U-Net. DO-U-Net is an Encoder-Decoder style, Fully Convolutional Network (FCN) for object segmentation and counting in image processing. Our proposed architecture achieves precision and sensitivity superior to other, similar models by producing two target outputs: a segmentation mask and an edge mask. Two case studies are used to demonstrate the capabilities of DO-U-Net: locating and counting Internally Displaced People (IDP) tents in satellite imagery, and the segmentation and counting of erythrocytes in blood smears. The model was demonstrated to work with a relatively small training dataset, achieving a sensitivity of 98.69% for IDP camps of the fixed resolution, and 94.66% for a scale-invariant IDP model. DO-U-Net achieved a sensitivity of 99.07% on the erythrocytes dataset. DO-U-Net has a reduced memory footprint, allowing for training and deployment on a machine with a lower to mid-range GPU, making it accessible to a wider audience, including non-governmental organisations (NGOs) providing humanitarian aid, as well as health care organisations.

Keywords: Convolutional neural networks · U-Net · Segmentation · Counting · Satellite imagery · Blood smear

1 Introduction

Over recent years, the volumes of data collected across all industries globally have grown dramatically. As a result, we find ourselves in an ever greater need for fully automated analysis techniques. The most common approaches to large scale data analysis rely on the use of supervised and unsupervised Machine Learning, and,

© The Author(s) 2020
M. R. Berthold et al. (Eds.): IDA 2020, LNCS 12080, pp. 391–403, 2020.
https://doi.org/10.1007/978-3-030-44584-3_31

increasingly, Deep Learning. Using only a small number of human-annotated data samples, we can train models to rapidly analyse vast quantities of data without sacrificing the quality or accuracy compared with a human analyst. In this paper, we focus on images - a rich datatype that often requires rapid and accurate analysis, despite its volumes and complexity. Object classification is one of the most common types of analysis undertaken. In many cases, a further step may be required in which the classified and segmented objects of interest need to be accurately counted. While easily performed by humans, albeit slowly, this task is often non-trivial in Computer Vision, especially in the cases where the objects exist in complex environments or when objects are closely co-located and overlapping. We look at two such case studies: locating and counting Internally Displaced People (IDP) shelters in Western Afghanistan using satellite imagery, and the segmentation and counting of erythrocytes in blood smear images. Both applications have a high impact in the real world and are in a need of new rapid and accurate analysis techniques.

1.1 Searching for Shelter: Locating IDP Tents in Satellite Imagery

Over 40 million individuals were believed to have been internally displaced globally in 2018 [1]. Afghanistan is home to 2,598,000 IDPs displaced by conflict and violence, with the numbers growing by 372,000 in the year 2018 alone. In the same year, an additional 435,000 individuals were displaced due to natural disasters, 371,000 of whom were displaced due to drought.

The internally displaced population receive aid from various non-governmental organisations (NGOs), to prevent IDPs from becoming refugees. The Norwegian Refugee Council (NRC) is one such body, providing humanitarian aid to IDPs across 31 countries, assisting 8.5 million people in 2018 [2]. In Afghanistan, the NRC has been providing IDPs with tents as temporary living spaces. Alcis is assisting the NRC with the analysis of the number, flow, and concentration of these humanitarian shelters, enabling valuable aid to be delivered more effectively.

Existing Methods. In the past, Geographical Information System (GIS) Technicians relied mostly on industry-standard software, such as ArcGIS, for the majority of their analysis. These tools provide the user with a small number of built-in Machine Learning algorithms, such as the popularly used implementation of the Support Vector Machine (SVM) algorithm [3]. These generally involve a time consuming, semi-automated process, with each step being revisited multiple times as manual tuning of the model parameters is required. The methodology does not allow for batch processing as all stages must be repeated with human input for each image. An example of the ArcGIS process[1] used by GIS technicians is:

1. Manually segment the image by selecting a sample of objects exhibiting similar shape, spectral and spatial characteristics.

[1] Details of the process have been provided by Alcis.

2. Train the image classifier to identify other examples similar to the marked sample, using a built-in machine learning model (e.g. SVM).
3. Identify any misclassified objects and repeat the above steps.

More recently, many GIS specialists have begun to look towards the latest techniques in Data Science and Big Data analysis to create custom Machine Learning solutions. A review paper by Quinn et al. in 2018 [4] weighed up the merits of Machine Learning approaches used to segment and count both refugee and IDP camps. Their work used a sample of 87,123 structures; a magnitude which was required for training using their approach and was seen as a limitation. Quinn et al. used the popular Mask R-CNN [5] segmentation model to analyse their data; a model using a Region Proposal Network to simultaneously classify and segment images. This yielded an average precision of 75%, improving to 78% by applying a transfer learning approach.

1.2 Counting in Vein: Finding Erythrocytes in Blood Smears

The counting of erythrocytes, or red blood cells, in blood smear images, is another application in which one must count complex objects. This is an important task in exploratory and diagnostic medicine, as well as medical research. An average blood smear imaged using a microscope, contains several hundred erythrocytes of varying size, many of which are overlapping, making an accurate manual count both difficult and time-consuming.

Existing Methods. While only a small number of analyses were able to successfully perform an erythrocyte count, Tran et al. [6] have achieved a counting accuracy of 93.30%. Their technique relied on locating the cells using the SegNet [7] network. SegNet is an encoder-decoder style FCN architecture producing segmentation masks as its output. Due to the overlap of erythrocyte cells, they performed a Euclidean Distance Transform on the binary segmentation masks to obtain the location of each cell using a connected component labelling algorithm. A work by Alam and Islam [8] proposes an approach using YOLO9000 [9]; a network using a similar approach to Mask R-CNN, to locate elliptical bounding regions that roughly correspond to the outer contours of the cells. Using 300 images, each containing a small number of erythrocytes, for training, they achieve an accuracy of 96.09%. Bounding boxes acted as ground-truth for Alam and Islam, as opposed to segmentation masks used by Tran et al.

2 Data Description

2.1 Satellite Imagery

Working on the ground, the NRC identified areas within Western Afghanistan with known locations of IDP camps. Through their relationship with Maxar

[10], Alcis has access to satellite imagery covering multiple camps, in a range of different environments. Figure 1 shows a section of a camp in Qala'e'Naw, Badghis.

This work uses imagery collected by the WorldView-2 and WorldView-3 satellites [11], by their operator and owner Maxar. WorldView-2 has a multispectral resolution of 1.85 m, while the multispectral resolution of WorldView-3 is 1.24 m [12], allowing tents of approximately 7.5 m long and 4 m wide to be resolved. The WorldView-2 images were captured on either 05/01/19 (DD/MM/YY) or 03/03/19, with the WorldView-3 images captured on 12/03/19. A further set of images, observed between 08/08/18 and 23/09/19 by WorldView-3, became available for some locations. This dataset can be used to show evolution in the camps during this period, allowing for a better understanding of the changes undergone in IDP camps. Due to the orbital position of the satellite, images observed at different times have varying resolutions, as well as other properties, due to differences in viewing angle and atmospheric effects.

Training Data. We developed DO-U-Net using a limited number of satellite images, obtained over a very limited time, with a nearly identical pixel scale. Each tent found in the training imagery has been marked with a polygon, using a custom Graphical User Interface (GUI) tool developed by Alcis. This has been done for a total of 6 images, covering an area of approximately 15 km^2 and containing 5,178 tents. Incidentally, this makes our training dataset nearly 17 times smaller than that used by Quinn et al. in their analysis.

The second satellite dataset includes imagery of varying quality and resolution, providing an opportunity to develop a scale-invariant version of our model. We used 3 additional training images, distinct from the original dataset, to train our scale-invariant DO-U-Net. These images contained 2,338 additional tents, in an area of around 130 km^2, giving a total of 7,516 tents in over 140 km^2.

2.2 Blood Smear Images

We used blood smear images from the Acute Lymphoblastic Leukemia (ALL) Image Database for Image Processing[2]. These images were captured using an optical laboratory microscope, with magnification ranging from 300–500×, and a Canon PowerShot G5 camera. We used the ALL_IDB1 dataset, comprised of 108 images taken during September 2005 from both ALL and non-ALL patients. An example blood smear image from an ALL patient can be seen in Fig. 2.

Training Data. We selected 10 images from the ALL_IDB1 dataset to be used as training data. These images are representative of the diverse nature of the entire dataset, including the varying microscope magnifications and backgrounds. Of the images used, 3 belong to ALL patients, with the remaining 7

[2] Provided by the Department of Information Technology at Università degli Studi di Milano, https://homes.di.unimi.it/scotti/all/.

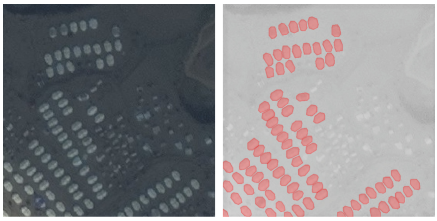

Fig. 1. *Left:* An IDP camp in Badghis, Afghanistan. NRC tents are clearly visible due to their uniform size and light colour. *Right:* The manually created ground-truth annotation for the image.

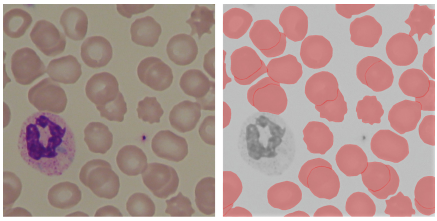

Fig. 2. *Left:* An image of a blood smear from an Acute Lymphoblastic Leukemia (ALL) patient. *Right:* The manually created ground-truth annotation for the image. The images also contain lymphocytes, which are not marked in the training data.

images coming from non-ALL patients. Similarly to the IDP camp dataset, all erythrocytes in the training data have been manually labelled with a polygon using our custom GUI tool. In the images belonging to ALL patients, a total of 1,300 erythrocytes were marked. A further 3,060 erythrocytes were marked in the images belonging to non-ALL patients, giving a total of 4,360 erythrocytes in the training data.

The training data does not distinguish between typical erythrocytes and those with any forms of morphology - of the 4,360 erythrocytes, just 106 display a clear morphology. The training data also does not contain any annotation for leukocytes. Instead, our focus is on correctly segmenting and counting all erythrocytes in the images.

3 Methodology

Of late, several very advanced and powerful Computer Vision algorithms have been developed, including the popular Mask R-CNN [5] and YOLO [9] architectures. While their performance is undoubtedly impressive, they rely on a large number of images to train their complex networks, as highlighted by Quinn et al. [5]. More recently, many more examples of FCN have been developed, including SegNet [7], DeconvNet [13] and U-Net [14], with the latter emerging as arguably one of the most popular encoder-decoder based architectures. Aimed at achieving a high degree of success even with sparse training datasets and developed to tackle biological image segmentation problems, it is a clear starting block for our architecture.

3.1 U-Net

The classical U-Net, as proposed by Ronneberger et al. has revolutionised the field of biomedical image segmentation. Similarly to other encoder-decoder networks, U-Net is capable of producing highly precise segmentation masks.

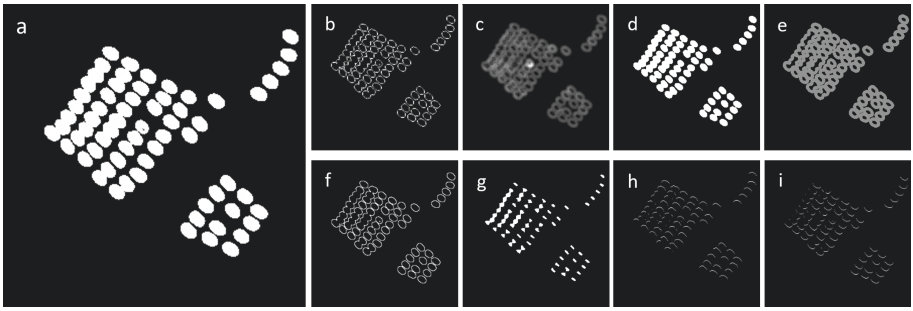

Fig. 3. *a*: Sample segmentation mask in which some tent segmentations are seen to overlap. *b*: Sobel filtered image. *c*: Local entropy filtered image. *d*: Otsu filtered image. *e*: Image with contour ellipses applied. *f*: Image with gradient morphology applied. *g*: Eroded image. *h*: Tophat filtered image. *i*: Blackhat filtered image.

What differentiates it from Mask R-CNN, SegNet and other similar networks is its lack of reliance on large datasets [14]. This is achieved by the introduction of a large number of skip connections, which reintroduce some of the early encoder layers into the much deeper decoder layers. This greatly enriches the information received by the decoder part of the network, and hence reduces the overall size of the dataset required to train the network.

We have deployed the original U-Net on our dataset of satellite images of IDP camps in Western Afghanistan. While we were able to produce segmentation masks that very accurately marked the location of the tents, the segmentation masks contained significant overlaps between tents, as seen in Fig. 3. This overlap prevents us from carrying out an automated count, despite using several post-processing techniques to minimise the impact of these overlaps. The most successful post-processing approaches are shown in Fig. 3. The issues encountered with the classical U-Net have motivated our modifications to the architecture, as described in this work.

3.2 DO-U-Net

Driven by the need to reduce overlap in segmentation masks, we modified the U-Net architecture to produce dual outputs, thus developing the DO-U-Net. The idea of a contour aware network was first demonstrated by the DCAN architecture [15]. Based on a simple FCN, DCAN was trained to use the outer contours of the areas of interest to guide the training of the segmentation masks. This led to improved semantic and instance segmentation of the model, which in their case, looked at non-overlapping features in biomedical imaging.

With the aim of counting closely co-located and overlapping objects, we are predominantly interested in the correct detection of individual objects as opposed to the exact precision of the segmentation mask itself. An examination of the hidden convolutional layers of the classical U-Net showed that the penultimate layer of the network extracts information about the edges of our objects of

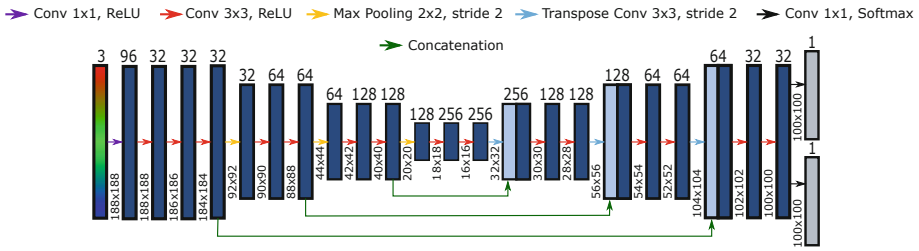

Fig. 4. The DO-U-Net architecture, showing two output layers that target the segmentation and edge masks corresponding to the training images.

interest, without any external stimulus. We introduce a secondary output layer to the network, targeting a mask segmenting the edges of our objects. By subtracting this "edge" mask from the original segmentation mask, we can obtain a "reduced" mask containing only non-overlapping objects.

As our objective was to identify tents of fixed scale in our image dataset, we were able to simplify the model considerably. This reduced the computational requirements in training of the model, allowing not only for much faster development and training but also opening the possibility of deploying the algorithm on a dataset covering a large proportion of the total area of Afganistan, driven by our commercial requirements.

Architecture. Starting with the classical U-Net, we reduce the number of convolutional layers and skip connections in the model. Simultaneously, we minimised the complexity of the model by looking at smaller input regions of the images, thus minimising the memory footprint of the model. We follow the approach of Ronneberger et. al. by using unpadded convolutions throughout the network, resulting in a model with smaller output masks (100×100 px) corresponding to a central region of a larger (188×188 px) input image region. DO-U-Net uses two, independently trained, output layers of identical size. Figure 4 shows our proposed DO-U-Net architecture. The model can also be found in our public online repository[3]. Examples of the output edge and segmentation masks, as well as the final "reduced" mask, can be seen in Figs. 6 and 7. With the reduced memory footprint of our model, we can produce a "reduced" segmentation mask for a single 100×100 px region in 3 ms using `TensorFlow 2.0` with Intel i9-9820X CPU and a single NVIDIA RTX 2080 Ti GPU setup.

Training. The large training images were divided such that no overlap exists between the regions corresponding to the target masks, using zero-padding at the image borders. We train our model against both segmentation and edge masks. The edges of the mark-up polygons, annotated using our custom tool, were used as the "edge" masks during training. Due to the difference in a pixel

[3] https://github.com/ToyahJade/DO-U-Net.

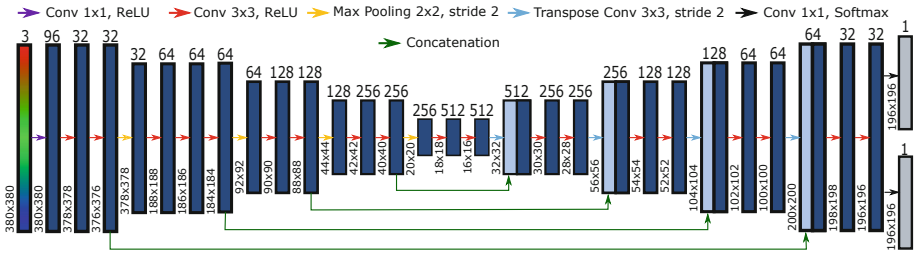

Fig. 5. Scale-Invariant DO-U-Net, redesigned to work with datasets containing objects of variable scale.

size of tents and erythrocytes, the edges were taken to span 2 px and 4 px wide respectively in these case studies.

As our problem deals with segmentation masks covering only a small proportion of the image (<1% in some satellite imagery), the choice of a loss function was a very important factor. We use the Focal Tversky Index, which is suited for training with sparse positive labels compared to the overall area of the training data [16]. Our best result, obtained using the Focal Tversky loss function, gave an improvement of 5% in the Intersect-over-Union (IoU) value compared to the Binary Cross-Entropy loss function, as used by Ronneberger et al. [14]. We found the training to behave most optimally when the combined loss function for the model was heavily weighted toward the edge mask segmentation. Here, we used a 10%/90% split for the object and edge mask segmentation respectively.

Counting. As the resulting "reduced" masks produced by our approach do not contain any overlaps, we can use simple counting techniques, relying on the detection of the bounding polygons for the objects of interest. We apply a threshold to remove all negative values from the image, which may occur due to the subtractions. We then use the Marching Squares Algorithm implemented as part of `Python`'s `skimage.measure` image analysis library [17].

Scale-Invariant DO-U-Net. In addition to the simple DO-U-Net, we propose a scale-invariant version of the architecture with an additional encoder and decoder block. Figure 5 shows the increased depth of the network as is required to capture the generalised model of our objects in the scale varying dataset. The addition of extra blocks resulted in a larger input layer of 380×380 px, corresponding to a segmentation mask of 196×196 px.

4 Results

4.1 IDP Tent Results

Using our DO-U-Net architecture, we were able to achieve a very significant improvement in the counting of IDP tents compared to the popularly used SVM

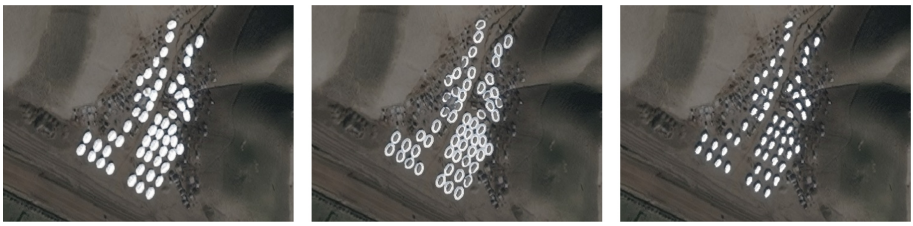

Fig. 6. *Left:* Segmentation mask produced for NRC tents in a camp near Qala'e'Naw. *Centre:* Edges mask produced for the same image. *Right:* The final output mask.

classifier available in ArcGIS. However, due to the manually intensive nature of the ArcGIS approach[4], we were only able to directly compare a single test camp, located in the Qala'e'Naw region of the Badghis Province. This area contains 921 tents as identified in the ground-truth masks. Using DO-U-Net, we achieved a precision of 99.78% with a sensitivity of 99.46%. Using ArcGIS, we find a precision of 99.86% and a significantly lower sensitivity of 79.37%. Sensitivity, or the true positive rate, measures the probability of detecting an object and is, therefore, the most important metric for us as we aim to locate and count all tents in the region. The scale-invariant DO-U-Net achieved a precision of 98.48% and a sensitivity of 98.37% on the same image.

We also apply DO-U-Net to a larger dataset of five images containing a total of 3,447 tents and find an average precision of 97.01% and an average sensitivity of 98.68%. Similarly, we tested the scale-invariant DO-U-Net using 10 images with varying properties and resolutions containing a total of 5,643 tents. Here, the average precision was reduced to 91.45%, and the average sensitivity dropped to 94.66%. This result is not surprising as, on inspection, we find that without the scale constraints the resulting segmenting masks are contaminated with other structures of similar properties to NRC tents. We also find that, without scale constraints, NRC tents which are partially covered e.g. with tarpaulin may be missed or only partially segmented. Our DO-U-Net and scale-invariant DO-U-Net sensitivities of 98.68% and 94.66% respectively are very strong results when compared to the existing literature.

4.2 Erythrocyte Results

To validate the performance of DO-U-Net at counting erythrocytes, we use 3 randomly selected blood smear images from ALL patients and a further 5 selected images from non-ALL patients. While randomly selected, the images are representative of the entire ALL_IDB1 dataset. On a total of 2,775 erythrocytes, as found in these 8 validation images, DO-U-Net achieved an average precision of 98.31% and an average sensitivity of 99.07%.

[4] Results found using ArcGIS methodology can be found at https://storymaps.arcgis.com/stories/d85e5cca27464d97ad4c1bad3da7f140.

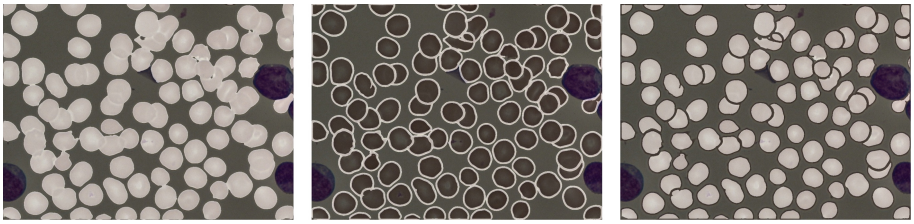

Fig. 7. *Left:* Segmentation mask produced for a blood smear of an ALL patient. *Centre:* Edges mask produced for the same image. *Right:* The final output mask.

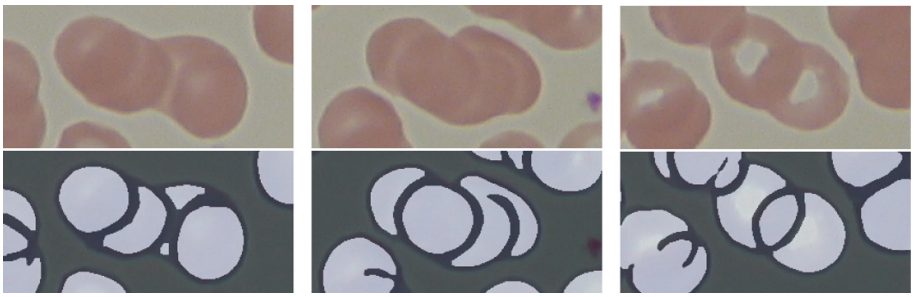

Fig. 8. *Top:* Blood smear images of overlapping cells. *Bottom:* Segmentation masks produced by DO-U-Net. *Left:* An overlapped cell is counted twice when the "edges" from neighbouring cells overlap and break up the cell. *Centre:* A cell is missed due to an incomplete edge mask. *Right:* An uncertainty in the order of the cell overlap leads to the intersect between two cells being counted as an additional cell.

Whilst our proposed DO-U-Net is extremely effective at producing image and edge segmentation masks, as demonstrated in Fig. 7, we do note that the obtained erythrocyte count may not always match the near-perfect segmentation. Figure 8 shows examples of the three most common issues found to occur in the final "reduced" masks. These mistakes arise largely due to the translucent nature of erythrocytes and the difficulty in differentiating between a cell which is overlapping another and a cell which is overlapped. While these cases are rare, this demonstrates that further improvements can be made to the architecture.

4.3 Future Work

Our current model has been designed to segment only one type of object, which is a clear limitation of our solution. As an example, the blood smear images from the ALL-IDP1 dataset contain normal erythrocytes as well as two clear types of morphology: burr cells and dacryocytes. These morphologies may be signs of disease in patients, though burr cells are common artefacts, especially known to occur when the blood sample is aged. It is therefore important to not only count all erythrocytes, but to also differentiate between their various morphologies.

Table 1. Summary of results for DO-U-Net, when tested on our two satellite imagery datasets and the ALL_IDB1 dataset.

Dataset	Number of images	Number of objects	Average precision	Average sensitivity
IDP Camps (Fixed Scale)	5	3,447	97.01%	98.69%
IDP Camps (Scale-Invariant)	10	5,643	91.45%	94.66%
ALL_IDB1	8	2,775	98.31%	99.07%

While our general theory can be applied to identifying different types of object, further modifications to our proposed DO-U-Net would be required.

5 Conclusion

We have proposed a new approach to segmenting and counting closely co-located and overlapping objects in complex image datasets. For this, we developed DO-U-Net: a modified U-Net based architecture, designed to produce both a segmentation and an "edge" mask. By subtracting the latter from the former, we can locate and spatially separate objects of interest before automatically counting them. Our methodology was successful on both of our case studies: locating and counting IDP tents in satellite imagery, and the segmentation and counting of erythrocytes in blood smear images. In the first case study, DO-U-Net increased our sensitivity by approximately 20% compared to a popular ArcGIS based solution, achieving an average sensitivity of 98.69% for a dataset of fixed spatial resolution. Our network went on to achieve a precision of 91.45% and a sensitivity of 94.66% on a set of satellite images with a varying resolution and colour profiles. This is an impressive result when compared to Quinn et al. who achieved a precision of 78%. We also found DO-U-Net to be extremely successful at segmenting and counting erythrocytes in blood smear images, achieving a sensitivity of 99.07% for our test dataset. This is an improvement of 6% over the results found by Tran et al. who used the same training dataset, and 3% over Alam and Islam who used a comparable set of images, giving us a near-perfect sensitivity when counting erythrocytes. The results are summarised in Table 1.

Acknowledgements. We thank Harry Robson, GIS Analyst at Alcis Holdings Ltd., for industry knowledge shared and for performing post-processing in ArcGIS. We would also like to show our gratitude to Richard Brittan, Tim Buckley and the Alcis team for sharing insight with us during the course of this research.

We are also immensely grateful to the Department of Information Technology at Università degli Studi di Milano for providing the ALL_IDB1 dataset from the Acute Lymphoblastic Leukemia Image Database for Image Processing.

References

1. Global Internal Displacement Database, Internal Displacement Monitoring Centre. http://www.internal-displacement.org/database/displacement-data. Accessed 04 Nov 2019
2. We Assisted 8.5 Million People Last Year, Norwegian Refugee Council. https://www.nrc.no/perspectives/2019/we-assisted-8.5-million-people-last-year. Accessed 04 Nov 2019
3. Train Support Vector Machine Classifier, ArcGIS Pro. https://pro.arcgis.com/en/pro-app/tool-reference/spatial-analyst/train-support-vector-machine-classifier.htm. Accessed 10 Nov 2019
4. Quinn, J.A., Nyhan, M.M., Navarro, C., Coluccia, D., Bromley, L., Luengo-Oroz, M.: Humanitarian applications of machine learning with remote-sensing data: review and case study in refugee settlement mapping. Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. **376**, 20170363 (2018). https://doi.org/10.1098/rsta.2017.0363
5. He, K., Gkioxari, G., Dollar, P., Girshick, R.: Mask R-CNN. Facebook AI research (FAIR). Submission Date: 24 January 2018. arXiv:1703.06870
6. Tran, T., Binh Minh, L., Lee, S., Kwon, K.: Blood cell count using deep learning semantic segmentation. Preprints 2019, 2019090075. https://doi.org/10.20944/preprints201909.0075.v1
7. Badrinarayanan, V., Kendall, A., Cipolla, R.: SegNet: a deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **39**(12), 2481–2495 (2017). https://doi.org/10.1109/TPAMI.2016.2644615. Dickinson, S. (ed.)
8. Alam, M.M., Islam, M.T.: Machine learning approach of automatic identification and counting of blood cells. Healthc. Technol. Lett. **6**(4), 103–108 (2019). https://doi.org/10.1049/htl.2018.5098
9. Redmon, J., Farhadi, A.: YOLO9000: better, faster, stronger, Submission Date: 25 December 2016. arXiv: 612.08242
10. Maxar Technologies. https://www.maxar.com/. Accessed 18 Nov 2019
11. WorldView-3, DigitalGlobe Inc. http://worldview3.digitalglobe.com. Accessed 04 Nov 2019
12. The DigitalGlobe Constellation, DigitalGlobe Inc. https://www.digitalglobe.com/company/about-us. Accessed 04 Nov 2019
13. Hyeonwoo, N., Seunghoon, H., Bohyung, H.: Learning deconvolution network for semantic segmentation. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1520–1528, May 2015
14. Ronneberger, Olaf, Fischer, Philipp, Brox, Thomas: U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, Nassir, Hornegger, Joachim, Wells, William M., Frangi, Alejandro F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
15. Chen, H., Qi, X., Yu, L., Heng, P.: DCAN: deep contour-aware networks for accurate gland segmentation. In: Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2487–2496. IEEE (2016)
16. Abraham, N., Khan, N.M.: A novel focal tversky loss function with improved attention U-Net for lesion segmentation. In: IEEE International Symposium on Biomedical Imaging, ISBI (2019). arXiv:1810.07842
17. Contour Finding, Scikits-image. https://scikit-image.org/docs/0.5/auto_examples/plot_contours.html. Accessed 11 Nov 2019

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.